Target Identification Based on the Transferable Belief Model Interpretation of Dempster-Shafer Model.

François Delmotte
LAMIH, Université de Valenciennes
Le Mont Houy, 59313 Valenciennes Cedex 9, FRANCE
francois.delmotte@univ-valenciennes.fr
and
Philippe Smets
IRIDIA - Université Libre de Bruxelles
50 av. Roosevelt, CP 194-6, 1050 Bruxelles, Belgium
psmets@ulb.ac.be
http://iridia.ulb.ac.be/˜psmets

Abstract

This paper explains how multisensor data fusion and target identification can be performed within the transferable belief model, a model for the representation of quantified uncertainty based on belief functions. We present the underlying theory, in particular the General Bayesian Theorem needed to transform likelihoods into beliefs and the pignistic transformation needed to build the probability measure required for decision making. We present how the this method applies in practice. We compare its solution with the classical one, illustrating it with an embarrassing example where the TBM and the probability solutions completely disagree. Computational efficiency of the belief function solution was supposedly proved in a study that we reproduce and we show that in fact the opposite conclusions hold. The results presented here can be extended directly to many problems of data fusion and diagnosis.

Keywords: Belief functions, transferable belief model, General Bayesian Theorem, pignistic probabilities, target identification, data fusion.

I. Introduction

Classically, uncertainty is represented by probability functions, but other models like those based on belief functions have been proposed to represent quantified uncertainty. Among them, the transferable belief model (TBM) is a model to represent quantified uncertainties based on belief functions and unrelated to any underlying probability model.

In this paper, we focus on the application of the TBM to classification problems. We focus on target identification problems, considering them as prototypical examples of object classification and pattern recognition. We show how to apply the TBM to such a problem. We explain what are the General Bayesian Theorem (GBT) and the pignistic probabilities, the two major tools used for classification. We compare the TBM solution to the probability solution. We consider also the issue of computational efficiency.

The TBM corresponds to an interpretation of the model initially developed by Shafer in his book [21]. The TBM has been presented in [31], (see also [29] for a recent survey). It represents weighted opinions, called here beliefs. The axiomatic justification of the model can be found in [28].

Several authors have already applied belief functions to classification problems, but usually they base their work only on Shafer's book, seemingly ignoring all later developments. For their defense, it must be acknowledged that the theoretical material about the TBM is scattered over many papers, some difficult to access. So we feel it useful for the reader to regroup the useful results here. Details and proofs can be found in the original papers.

The computational efficiency of the belief function approach and its comparison with the probability approach has been studied in [32], [4] on a problem of target identification. The comparison is based on the number of identical data that must be collected in order to reach a 'firm' decision. Unfortunately, these authors use belief functions in a way we feel to be ad hoc. These authors conclude that the probability approach is more efficient. We present in details how we consider belief functions should be applied to such problems. We repeat the efficiency study and show, on the contrary, that the TBM approach is computationally *more* efficient than the probability one.

This paper is organized as follows. In Section II, we present the general problem of target identification by multisensors and the way the TBM can be applied to this problem. In Section III, we present the General Bayesian Theorem. In Section IV, we apply the General Bayesian Theorem to a simple problem. In Section V, we show through an example that the TBM and the probability solutions can be diametrically opposed indicating thus that the choice of the model can be an essential issue. In Section VI, we reproduce part of the study presented in [32], [4] and show that the TBM approach is computationally more efficient than the probability approach. In Sections VII, VIII and IX, we show how to use the TBM for multisensor target identification problems. In particular, we show that the TBM solution can produce answers different from those reached with the classical likelihood based methods. We conclude in Section X.

II. The Transferable Belief Model for Identification

A. Sensors and Identification

A sensor can be seen as a piece of equipment that observes some data $x \subseteq X$ and transmits some 'opinion' about the actual value of a parameter of interest $h \subseteq H$. In probability theory, the relation between X and H is represented by a probability distribution on X for each $h_i \in H$. Let $P(.|h_i)$ denote the probability measure on X given $h_i \in H$, where X and H are called the observation and hypothesis domains, respectively. After observing $x \subseteq X$, the sensor communicates its opinion on the value of H under the form of a 'likelihood' vector. Let $l(h_i|x)$ denote the likelihood of the hypothesis h_i given the observed data is x: by definition $l(h_i|x) = P(x|h_i)$. Inference on H is based on $l(h_i|x)$ and some *a priori* probabilities.

The rule of Bayes enable us to update our knowledge, and give us the posterior probabilities:

$$P(h_i|x) = \frac{l(h_i|x)P(h_i)}{\sum_{h_j \subseteq H} l(h_j|x)P(h_j)}$$

The TBM enables us to introduce more complex types of uncertainty, at both the likelihoods and the *a priori* level.

B. The Transferable Belief Model

The idea of using belief functions to represent quantified uncertainty was first introduced by Shafer [21] who was building his theory from Dempster's research [7]. Later Gordon and Shortliffe [10] coined the term 'Dempster-Shafer's theory'. Unfortunately, this term turns out to be ambiguous as it does not differentiate between a model based on interval valued probability functions (not considered here) and the TBM. Missing to distinguish between these models has created confusion in the literature [27].

The central element of the TBM is the basic belief assignment (bba), denoted m. For $A \subseteq H$, $m(A)$ is the part of belief that supports A (i.e. the fact that the actual value h_0 of H is in A), and that, due to a lack of information, does not support any strict subset of A. The initial total belief is scaled to 1, and thus $m(A) \in [0, 1]$, with $\sum_{A \subseteq H} m(A) = 1$. We do not require $m(\emptyset) = 0$ as in Shafer's work.

The degree of belief $bel(A)$ is defined as: $bel : 2^H \rightarrow [0, 1]$ with, for all $A \subseteq H$,

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \tag{1}$$

where 2^H means all the subsets of the set H.

It quantifies the total amount of 'justified specific' support given to A. The term 'justified' means that B supports A, thus $B \subseteq A$, and the term 'specific' means that B does not support $\overline{A}$, thus $B \not\subseteq \overline{A}$ or equivalently $B \neq \emptyset$.

The degree of plausibility $pl(A)$ is defined as: $pl : 2^H \rightarrow [0, 1]$ with, for all $A \subseteq H$,

$$pl(A) = \sum_{B \subseteq X, B \cap A \neq \emptyset} m(B) = bel(H) - bel(\overline{A}). \tag{2}$$

It quantifies the maximum amount of 'potential specific' support that could be given to A. The term 'potential' means that B might come to support A without supporting $\overline{A}$ if further piece of evidence is taken into consideration, thus $B \cap A \neq \emptyset$.

The commonality function q is defined as: $q : 2^H \rightarrow [0, 1]$ with, for all $A \subseteq H$,

$$q(A) = \sum_{B \subseteq X, : A \subseteq B} m(B). \tag{3}$$

The functions m, bel, pl and q are always in one to one correspondence. They all describe the same information but seen under different points of view.

C. Notation

In order to enhance the fact that we work with non-normalized belief functions ($m(\emptyset)$ may be positive), we use the notation *bel* and *pl*, whereas Shafer uses the notation Bel and Pl. The latter two are kept for normalized belief and plausibility functions, i.e., where $m(\emptyset) = 0$.

Besides we use the next conventions. The notation for *bel* and its related functions is:

$$bel^H[F](A) = x.$$

It denotes that x is the value of the degree of belief that the actual value h_0 of H belongs to the set A, where A is a subset of the frame of discernment H. The belief is based on the facts F.

In the above notation, *bel* can be replaced by any of m, *pl*, q, etc... The indices should made it clear what the links are.

Note that $bel^H[F]$ denotes the belief function, and can be understood as a finite vector of length $2^{|H|}$, which components are the values of $bel^H[F](A)$ for every $A \subseteq H$.

We also take the convention that h_i denotes an element of H whereas h denotes a subset of H.

D. Conjunctive Combinations

Let E_1 and E_2 be two 'distinct' pieces of evidence and let $m^H[E_1]$ and $m^H[E_2]$ be the bba's they induce on H. Remember the symbols between [and] denote the pieces of evidence taken in consideration when building the belief functions.

We want to build the bba $m^H[E_1, E_2] = m^H[E_1] \odot m^H[E_2]$ that results from the combination of the two pieces of evidence provided both sources are accepted as fully reliable.

The conjunctive combination $m^H[E_1] \odot m^H[E_2]$ is defined by:

$$m^H[E_1] \odot m^H[E_2](A) = \sum_{B,C \subseteq H, B \cap C = A} m^H[E_1](B) m^H[E_2](C), \ \forall A \subseteq H \tag{4}$$

Dempster's rule of combination is obtained by normalizing the result of the conjunctive combination rule, i.e., by dividing all results by $(1 - m^H[E_1] \odot m^H[E_2](\emptyset))$. The result is denoted then by $m^H[E_1] \oplus m^H[E_2]$.

The conjunctive combination can also be computed as:

$$q^H[E_1] \odot q^H[E_2](A) = q^H[E_1](A) q^H[E_2](A), \qquad \forall A \subseteq H. \tag{5}$$

E. Conditioning

A special case of conjunctive combination rule is the conditioning rule. Let m_A be so that $m_A(X) = 1$ if $X = A$, and 0 otherwise. The result of the conjunctive combination of m with m_A produces a new bba $m[A]$ with:

$$m[A](B) = \sum_{C \subseteq \overline{A}} m(B \cup C) \qquad \text{if } B \subseteq A \tag{6}$$

$$= 0 \qquad \text{otherwise} \tag{7}$$

$$bel[A](B) = bel(B \cup \overline{A}) - bel(\overline{A}), \qquad \forall B \subseteq H \tag{8}$$

$$pl[A](B) = pl(A \cap B), \qquad \forall B \subseteq H. \tag{9}$$

This operation represents the impact of the information '$h_0 \notin \overline{A}$', which differs from $h_0 \in A$ in the open world context as h_0 might then not belong in H.

F. The Least Commitment Principle

An hypothesis that receives a belief of .3 is less supported than an hypothesis that receives a belief of .4. If $bel_1(h) \leq bel_2(h)$ for every $h \subseteq H$, we can say that bel_1 gives less support to (every hypotheses h of) H than bel_2. We say that bel_1 is less committed than bel_2. For normalized belief functions, these inequalities are the same as: $pl_1(h) \geq pl_2(h)$ for all $h \subseteq H$.

With non normalized belief functions, the inequalities between the bel_i are not adequate, whereas those between the pl_i are still adequate [29]. The formal definition becomes:

Let two bba m_1 and m_2 be defined on a frame of discernment H. m_1 is less committed than m_2 iff: $pl_1(h) \geq pl_2(h),\ \forall h \subseteq H$.

The least committed belief function defined on H is the vacuous belief function defined by $pl(h) = 1$, $\forall h \subseteq H$, or equivalently $m(h) = 1$ if $h = H$, and 0 otherwise. It represents a state of total ignorance as none of the strict subsets of H is supported.

G. Extensions and Marginalizations

G.1 The Vacuous Extension.

Let H be a frame of discernment. Let H' be a refinement R of H, i.e., every element h_i of H is mapped by R into one to several elements of H', and the images $R(h_i)$ and $R(h_j)$ on H' of any pair $h_i, h_j \in H, h_i \neq h_j$ are disjoint. It just means that H' is more detailed than H. Suppose there is a bba m^H on H. This bba can be extended on H' in order to build a bba on H' that expresses the same information as contained in m^H. This transformation is called the vacuous extension of m^H on H', denoted by $m^{H\uparrow H'}$ and its values are given by:

$$\begin{aligned} m^{H\uparrow H'}(h') &= m^H(h) \qquad && \text{if } h' = R(h) \\ &= 0 && \text{otherwise.} \end{aligned}$$

where $R(h)$ is the image of h under R [21].

G.2 Coarsening.

Suppose a bba $m^{H'}$ defined on H'. Let H be a coarsening of H', i.e., H' is a refinement R of H. The bba induced on H by $m^{H'}$ is denoted by $m^{H'\downarrow H}$, and the values of its related $bel^{H'\downarrow H}$ are:

$$bel^{H'\downarrow H}(h) = bel^{H'}(R(h)) \qquad \forall h \subseteq H$$

Marginalization is a special case of coarsening when H can be represented as the product space of two variables X and Y, and the bba defined on $X \times Y$ is transformed into a bba on X; indeed X is a coarsening of $X \times Y$.

G.3 The Ballooning Extension.

Let H be a frame of discernment and let H' be a subset of H. Suppose a bba $m^{H'}$ defined on H' and we need a bba on H. The least committed bba on H such that its conditioning on H' is $m^{H'}$ is given by the 'ballooning' extension introduced in [22], [26], denoted $m^{H' \Uparrow H}$, and its values are:

$$\begin{aligned} m^{H' \Uparrow H}(h) &= m^{H'}(h') \qquad && \text{if } h' \subseteq H', h = h' \cup \overline{H'} \\ &= 0 && \text{otherwise.} \end{aligned}$$

The ballooning extension is useful when the received beliefs were build on a limited frame and we discover that some alternatives had not been taken into consideration when the sensor produced the bba on the limited frame. We can thus build a bba on the larger frame from the one collected on the limited frame. It is repeatedly used to derive the General Bayesian Theorem presented in Section III.

H. Credal and Pignistic Levels

In the TBM, beliefs held by an agent are represented by a belief function bel^H. When a decision must be taken by this agent and the optimal decision depends on the actual value of H, the belief function bel^H is transformed into a probability function on H, probability function which is used to compute the expected utilities required to select the optimal decision [20]. In order to enhance that this probability function does not represent the agent's beliefs, we call it the pignistic probability function and denote it $BetP^H$.

The transformation between the belief function and the pignistic probability function is called the pignistic transformation (see Section II-I).

Conceptually, the TBM distinguishes between two mental levels:

- the 'credal' level where beliefs are entertained and represented by belief functions and
- the 'pignistic' level where beliefs are used to make decisions and induce a pignistic probability function.

The qualifiers come from 'credo' I believe and 'pignus' a bet, both in Latin.

In probability theory, these two levels are not distinguished and probability functions quantify beliefs at both levels.

I. The Pignistic Probabilities for Decision Making

Suppose a bba m^H that quantifies beliefs on H. Its pignistic transformation producing the pignistic probability function $BetP^H$, is assumed to be a function of H and m^H,

$$BetP^H = \Gamma(m^H, H).$$

The only transformation from m^H to $BetP^H$ that satisfies some rationality requirements is the so called pignistic transformation given by:

$$BetP^H(h) = \sum_{A: h \in A \subseteq H} \frac{m^H(A)}{|A|(1 - m^H(\emptyset))}, \qquad \forall h \in H \tag{10}$$

where $|A|$ is the number of elements in A [25], [30], [31].

It is easy to show that the function $BetP^H$ is indeed a probability function and the pignistic transformation of a probability function is the probability function itself. We call it 'pignistic' in order to avoid the confusion that would consist in interpreting $BetP^H$ as a measure representing beliefs on H. $BetP$ is the probability function needed to determine the optimal decision, not the measure of some beliefs, the last being quantified by the belief function.

This approach has been shown to resist to the Dutch book argument used by the Bayesians to justify the probabilistic approach [31].

III. The General Bayesian Theorem

In probability theory, Bayes theorem permits the computation of a probability function over the space H given the value of some variable $x \in X$ from the knowledge of the probabilities over X given each $h_i \in H$, and some *a priori* probability function over H. The same idea has been extended in the TBM context where we will build a belief function over H given an observation $x \subseteq X$ from the knowledge of the belief function over X given each $h_i \in H$ and a vacuous *a priori* belief over H, i.e., an *a priori* describing a state of total ignorance (therefore solving the delicate problem of choosing an appropriate *a priori*). This generalization is called the General Bayesian Theorem or GBT for short.

Incorporating an *a priori* information on H is achieved by combining the belief induced by x on H computed with the GBT with the *a priori* belief using the conjunctive combination rule.

For the GBT, all that is needed from the sensor after it observes x is the vector of plausibilities $pl^X[h_i](x)$ for all $h_i \in H$. In many cases, the conditional belief over X given h_i is in fact represented by a probability function, in which case $pl^X[h_i](x) = P^X[h_i](x)$. In order to keep with the tradition, we call $pl^X[h_i](x)$ the likelihood of h_i given x, what we denote by $l(h_i|x)$.

Given the the likelihoods $l(h_i|x)$ for every $h_i \in H$, then for $x \subseteq X$ and for every $A \subseteq H$, Smets [22] proves:

$$m^H[x](A) = \prod_{h_i \in A} l(h_i|x) \prod_{h_i \in \overline{A}} (1 - l(h_i|x)) \tag{11}$$

$$bel^H[x](A) = \prod_{h_i \in \overline{A}} (1 - l(h_i|x)) - \prod_{h_i \in H} (1 - l(h_i|x)) \tag{12}$$

$$pl^H[x](A) = 1 - \prod_{h_i \in A} (1 - l(h_i|x)) \tag{13}$$

$$q^H[x](A) = \prod_{h_i \in A} l(h_i|x) \tag{14}$$

Should there be some non vacuous beliefs on H, represented by $m^H[E_0]$, then this belief is simply combined with $m^H[x]$ by the application of the conjunctive rule of combination.

The GBT has been derived axiomatically by Smets [22], [23], [26] and independently by Appriou [1].

A. Some Properties

Some particular properties of the GBT are worth mentioning.

A.1 A Way to Derive the GBT

The GBT can be derived from the ballooning extensions (see Section II-G). We start from the bba $m^X[h_i]$ collected for $h_i \in H$. We build its ballooning extension on $X \times H$, and then conjunctively combine these bba's over the h_i's. The result is then marginalized on H and is exactly the bba derived by the GBT.

Formally we have:

$$\begin{aligned} m^H[x] &= (\textcircled{\cap}_{h_i \in H}\ m^X[h_i]^{\Uparrow X\times H})[x]^{\downarrow H} \\ &= \textcircled{\cap}_{h_i \in H}\ (m^X[h_i]^{\Uparrow X\times H}[x]^{\downarrow H}) \end{aligned}$$

The combination can be performed before or after conditioning on x and marginalizing, results are identical.

Furthermore, the bba $m^X[h_i]^{\Uparrow X\times H}[x]^{\downarrow H}$ happens to be a simple support functions[1] on H with a mass $1 - l(h_i|x)$ given to $\overline{h_i}$ and a mass $l(h_i|x)$ given to H [26].

A.2 Independent Observations

Suppose two 'independent' observations x defined on X and y defined on Y, and the inference on H obtained from their joint observation.

Suppose the two variables X and Y are conditionally 'independent' given H. If the beliefs over X and over Y given $h_i \in H$ are represented by probability functions, it means that X and Y are conditionally stochastically independent given H. In the more general case where the beliefs over X and over Y given $h_i \in H$ are represented by belief functions, the 'independence' requirement becomes what is called the Conditional Cognitive Independence [26]. In both cases, the next property is satisfied;

$$l(h|x,y) = l(h|x)l(h|y)\ \forall h \in H, x \in X, y \in Y. \tag{15}$$

where the likelihood is either the conditional probability or the conditional plausibility of x given h.

The GBT could then be applied in two different ways.

Let $pl^H[x]$ and $pl^H[y]$ be computed by the GBT (with a vacuous *a priori* belief on H) from the likelihoods on H obtained from x and y, separately. They are then combined by the conjunctive rule of combination in order to build $pl^H[x,y]$.

We could as well consider the likelihoods directly obtained from the joint observation x, y, using the product rule (15). We then compute $pl^H[x,y]$ from them using the GBT.

Both results are the same. This property is essential and in fact at the core of the axiomatic derivations of the rule. Other suggested extensions of Bayes theorem to belief

[1] A Simple Support Function is a belief function where all masses are null except for one set and the whole frame.

functions fail to satisfy it.

A.3 New Hypothesis

The mass given to the empty set by the application of the GBT can receive a nice and useful interpretation. Suppose we accept that H is not exhaustive, thus there are unthought-of hypotheses. Let h^* denote all of them. Then $m^H[x](\emptyset)$ is equal to $bel^{H\cup h^*}[x](h^*)$, thus the degree of belief that the data x supports that none of the hypothesis in H holds, and that we are facing a case where a new previously unthought-of hypothesis must be considered. Identically we have for all $h \subseteq H$:

$$bel^{H\cup h^*}[x](h \cup h^*) = bel^H[x](h) + m^H[x](\emptyset),$$
$$bel^{H\cup h^*}[x](h) = 0.$$

This result is what the GBT produces if we add an extra hypothesis h^*, and define $bel^X[h^*]$ as the vacuous belief function. This is the natural solution, as it is obvious that the user knows nothing about the conditional beliefs over the data when the actual hypothesis belongs to h^*, the set of unthought-of hypotheses.

A.4 The Bayesian Degradation

If for each $h_i \in H$, $pl^X[h_i]$ is a probability function $P^X[h_i]$ on X, then the GBT for $|h_i| = 1$ becomes:

$$pl^H[x](h_i) = P(x|h_i), \qquad \forall x \subseteq X.$$

That is, on the singletons h_i of H, $pl^H[x]$ reduces to the likelihood of h_i given x. The analogy stops there as the solution for the likelihood of subsets of H are different.

If, furthermore, the *a priori* belief on H is also a probability function $p_0(h)$, then the normalized GBT becomes:

$$Bel^H[x](A) = \frac{\sum_{h_i\in A} P(x|h_i)p_0(h_i)}{\sum_{h_i\in H} P(x|h_i)p_0(h_i)} = P[x](A)$$

i.e. the (normalized) GBT reduces itself into the classical Bayesian theorem (as it should). This explains the origin of its name.

A.5 Most Plausible and Most Probable Hypothesis

Suppose we know $pl^X[h_i](x)$ for all $x \subseteq X$ and all $h_i \in H$. Data $x \subseteq X$ is collected and we want to select which hypothesis $h_i \in H$ is 'best supported' given the observed data. Two strategies have been proposed, one based on selecting the hypothesis with the largest pignistic probability (see Section II-I), the other based on selecting the hypothesis with the largest plausibility [1], [19] which is equivalent to selecting the hypothesis with the largest likelihood. It happens that the selected hypothesis is the same for both approaches when the *a priori* belief on H is vacuous.

Theorem III.1: Given $x \subseteq X$ and $pl^X[h_i](x)$ for all $h_i \in H$, let $pl^H[x]$ be the plausibility function defined on H and computed by the GBT (relations (11) to (14)), and $BetP^H[x]$ be the pignistic probability function constructed on H from $pl^H[x]$ (relation (10)), then:

$$BetP^H[x](h_i) > BetP^H[x](h_j) \qquad \text{iff} \qquad pl^X[h_i](x) > pl^X[h_j](x).$$

Proof. Let $l(h_k|x) = pl^X[h_k](x)$. Suppose $l(h_k|x) < 1$ for all $h_k \in H$. Let $r_k = l(h_k|x)/(1 - l(h_k|x))$ and $\alpha = \prod_{h_k \in H}(1 - l(h_k|x))$. Then by the GBT (see relation (11)), we have for $h \subseteq H$:

$$m^H[x](h) = \prod_{h_i \in h} l(h_i|x) \prod_{h_i \in \overline{h}} (1 - l(h_i|x)) = \alpha \prod_{h_k \in h} r_k.$$

With $K = 1/(1 - m^H[x](\emptyset))$, we have $BetP^H[x](h_i)$:

$$\begin{aligned} BetP^H[x](h_i) &= K \sum_{h \subseteq \overline{h_i}} \frac{1}{|h|+1} m^H[x](h_i \cup h) \\ &= \alpha K\ r_i \sum_{h \subseteq \overline{h_i}} \frac{1}{|h|+1} \prod_{h_k \in h} r_k \\ &= \alpha K \sum_{h \subseteq \overline{h_i \cup h_j}} \prod_{h_k \in h} r_k \ (\frac{r_i}{|h|+1} + \frac{r_i\ r_j}{|h|+2}) \text{ where } j \neq i \end{aligned}$$

In that case:

$$\begin{aligned} BetP^H[x](h_i) - BetP^H[x](h_j) &= \alpha\ K \sum_{h \subseteq \overline{h_i \cup h_j}} \prod_{h_k \in h} r_k \ (\frac{r_i}{|h|+1} - \frac{r_j}{|h|+1}) \\ &= \alpha\ K(r_i - r_j) \sum_{h \subseteq \overline{h_i \cup h_j}} \prod_{h_k \in h} r_k \ (\frac{1}{|h|+1}) \end{aligned}$$

As $r_k \geq 0$, the product terms are non negative, and so is their sum. The sum is positive as the term with $h = \emptyset$ has a product equal to 1. Hence the sign of the difference is the same as the sign of $r_i - r_j$:

$$BetP^H[x](h_i) > BetP^H[x](h_j) \text{ iff } r_i > r_j.$$

As $r_i > r_j$ iff $l(h_i|x) > l(h_j|x)$, i.e., iff $pl^X[h_i](x) > pl^X[h_j](x)$, the largest value of $BetP^H[x](h_\nu)$ is obtained for the hypothesis h_ν for which $pl^X[h_\nu](x)$ is maximal.

If for $k \in H_0 \subseteq H$, $l(h_k|x) = 1$, then every positive mass on H is given to a superset of h_0, and thus the pignistic probabilities given to the $h_k \in H_0$ are equal and maximal. Simultaneously, $pl^X[h_k](x)$ is always less or equal to 1, so the hypothesis $h_k \in H_0$ are those with a maximal plausibility, hence the theorem. □

This property is very useful when the only purpose is to take a decision and the *a priori* belief on H is vacuous. Indeed all computation can be avoided as all that is needed is $pl^X[h_k](x)$. Of course, the whole computation is still needed when expected utilities and other results are required. This result does not hold when a non vacuous *a priori* belief on H is introduced.

IV. A Simple Example of Data Fusion

In order to illustrate the use of the GBT and the pignistic transformation, we present a simple problem of target identification by two sensors. Our examples are inspired by those in [4].

Let S_1 and S_2 be two sensors, an Electronic Support Measure (ESM) and a Radar sensor, respectively. Let X and Y be the domains of the data they can observed, respectively. Let $H = \{F, M, B\}$ be the set of possible targets where the letters denote a F-15, a Mig-27 and a Boeing 747 aircraft, respectively. Table I presents the values of the conditional plausibility functions for $x \subseteq X$ and $y \subseteq Y$, where x and y are the observations made by the two sensors, respectively. Table II presents the computation performed by the GBT. We list the plausibility function induced by data x on H (line $pl^H[x]$), and its related bba (line $m^H[x]$) and commonality function (line $q^H[x]$) using relation (5). For instance,

$$pl^H[x](F, M) = 1 - (1 - .7)(1 - .4) = .82$$

We do the same with data y. We then conjunctively combine the two belief functions by a pointwise multiplication of their commonality functions (line $q^H[x,y]$). We then present the bba (line $m^H[x,y]$) and the normalized belief (line $Bel^H[x,y]$) and plausibility functions (line $Pl^H[x,y]$) related to the commonality function and that result thus from the conjunctive combination of the belief functions induced on H by x and by y. Finally Table III presents the pignistic probability functions computed after collecting data x alone, data y alone and data x and y jointly. For example,

$$BetP^H[x](F) = (.378 + .252/2 + .041/ - +.028/3)\ /\ (1 - .162) = .638$$

The ESM sensor supports the hypothesis that the object is a F-15, whereas the Radar sensor supports that it is a Boeing. But together, they support more strongly that the object is a F-15. This fits with what Table I tells. Hypothesis M and B are each rejected by one sensor, and F is the only hypothesis somehow supported by both sensors. For comparison purposes we also present in Table III what would be the posterior probabilities obtained with the same data using equi *a priori* probabilities on H. Results are very similar in this case. Such a similarity is not always encountered as shown in the nest example, where the TBM conclusions and the probability ones diverge completely.

Sensor	data	F	M	B
ESM	x	.7	.4	.1
Radar	y	.5	.2	.6

TABLE I

EXAMPLE 1. VALUES OF THE CONDITIONAL PLAUSIBILITY FUNCTIONS ON $x \subseteq X$ AND $y \subseteq Y$ GIVEN THE 3 HYPOTHESES F, M AND B IN H. x AND y ARE THE OBSERVATIONS MADE BY THE ESM AND THE RADAR SENSORS, RESPECTIVELY.

V. An Embarrassing Example Comparing the TBM with a Probability Approach

We present an example where the probability and the TBM approaches strongly disagree. This example is useful is showing that the choice between the two models can be

H	$\emptyset$	F	M	B	F, M	F, B	M, B	F, M, B
$pl^H[x]$	0	0.7	0.4	0.1	0.82	0.73	0.46	0.838
$m^H[x]$	0.162	0.378	0.108	0.018	0.252	0.041	0.012	0.028
$q^H[x]$	1	0.7	0.4	0.1	0.28	0.07	0.040	0.028
$pl^H[y]$	0	0.5	0.2	0.6	0.6	0.8	0.68	0.84
$m^H[y]$	0.16	0.16	0.039	0.24	0.04	0.24	0.060	0.06
$q^H[y]$	1	0.5	0.2	0.6	0.1	0.3	0.12	0.06
$q^H[x, y]$	1	0.35	0.08	0.06	0.028	0.021	0.004	0.001
$m^H[x, y]$	0.562	0.302	0.048	0.035	0.026	0.019	0.003	0.001
$Bel^H[x, y]$	0	0.691	0.111	0.081	0.862	0.817	0.200	1
$Pl^H[x, y]$	0	0.799	0.182	0.137	0.918	0.888	0.308	1

TABLE II

EXAMPLE 1. COMPUTATION PERFORMED BY THE GBT IN ORDER TO COMPUTE THE BELIEF AND PLAUSIBILITY FUNCTIONS ON H GIVEN THE x AND y DATA.

essential in practice. The example cannot be used to 'prove' that one of the two models is right or wrong, as neither common sense nor rationality requirements can definitively help us in deciding which of the two diverging conclusions is the 'correct' one. Some hints about this choice are discussed at the end of this section.

A. The Problem

Let $S_1, S_2 \dots S_{30}$ be a set of 30 sensors. Each sensor identifies objects as 'friend' (denoted Frd) or 'foe' (denoted Foe). For each sensor, we know what is the probability that it is in working condition or broken. Let $p_i = P(S_i \textit{ in Working Condition})$.

Let X_i be the domain of the data sensor S_i can collect, with $X_i = \{x_i, y_i\}$. When in working condition, the sensor is a perfect detector, i.e., x_i is equivalent to Frd and y_i to Foe. When broken, we totally ignore how the sensor would react when observing a friend or a foe.

All sensors are either of high quality (HQS) or of low quality (LQS), with sensors $S_1 \dots S_{11} \in HQS$ whereas $S_{12} \dots S_{30} \in LQS$. For each $S_i \in HQS$, $p_i = .99$ and for each

Sensor		F	M	B
ESM	$BetP^H[x]$	0.638	0.298	0.065
Radar	$BetP^H[y]$	0.381	0.131	0.488
ESM+Radar	$BetP^H[x, y]$	0.745	0.147	0.109
ESM+Radar	$P^H[x, y]$	0.714	0.163	0.122

TABLE III

EXAMPLE 1. PIGNISTIC PROBABILITIES COMPUTED ON H GIVEN OBSERVED DATA ARE x, y AND x, y, RESPECTIVELY. THE LAST LINE PRESENTS THE PROBABILITIES ONE WOULD OBTAINED WHEN APPLYING A CLASSICAL PROBABILITY APPROACH, USING AN EQUI *a priori* PROBABILITY FUNCTION ON H.

$S_i \in LQS,\ p_i = .90$.

The collected data is the vector $data = (x_1 \dots x_{10}, y_{11}, x_{12}, y_{13} \dots y_{30})$. So 10 out of the 11 high quality sensors and one among the 19 low quality sensors support the target is a friend, the others support it is a foe. What should we conclude after fusing these data? It is hard to decide as common sense can hardly help. We show now that the TBM concludes with a probability of .91 that the target is a friend, whereas the probability analysis concludes with a probability .92 that it is a foe. Values were of course chosen in order to get this enormous discrepancy. Nevertheless such a discrepancy is quite embarrassing. It shows that selecting the model deserves serious attention as conclusions can strongly depend on this choice.

B. Bayesian Analysis.

In order to proceed with a Bayesian analysis, we need first to assess $\alpha = P(x_i|S_i = Brk, Frd) = P(x_i|S_i = Brk, Foe)$, where we accept that the behavior of the broken sensor does not depend on the target's nature. A strict Bayesian claims that a probability can be assigned to any event, and thus that α can and must be assessed. The most natural assessment here (and the one most users will apply in practice) is $\alpha = .5$.

The Bayesian analysis proceeds then as follows. Table IV presents relevant data. We

Target	Frd			Foe		
S_i Status	Wrk	Brk		Wrk	Brk	
$P(Status)$	p_i	$1-p_i$		p_i	$1-p_i$	
X_i	$p(d_i)$	$p(d_i)$	$p(d_i\|Frd)$	$p(d_i)$	$p(d_i)$	$p(d_i\|Foe)$
x_i	1	.5	$p_i+.5(1-p_i)$	0	.5	$.5(1-p_i)$
y_i	0	.5	$.5(1-p_i)$	1	.5	$p_i+.5(1-p_i)$

TABLE IV

EXAMPLE 2. PROBABILITY APPROACH. FOR SENSOR S_i, VALUES OF $p(d_i)$ WHERE THE TARGET T CAN BE Frd (FOR FRIEND) OR Foe AND S_i STATUS CAN BE Wrk (FOR WORKING) AND Brk (FOR BROKEN). p_i IS THE PROBABILITY SENSOR S_i STATUS IS Wrk. THE COLUMNS $p(d_i|Frd)$ AND $p(d_i|Foe)$ PRESENT THE PROBABILITY ON X_i GIVEN THE TARGET IS Frd OR Foe, THUS THE LIKELIHOODS GIVEN TO THE TARGETS WHEN THE OBSERVED DATA IS x_i OR y_i, RESPECTIVELY.

must compute $P(Frd|data)$ where $data$ is the $data = (d_1 \dots d_{30})$ vector . We have:

$$\begin{aligned} P(Frd|data) &\propto P(Frd)P(data|Frd) \\ &= P(Frd) \prod_{i=1\dots30} P(d_i|Frd) \end{aligned}$$

assuming the conditional independence of the data given the nature of the target.

We have then, with Wrk_i and Brk_i being the status of S_i:

$$\begin{aligned} P(d_i|Frd) &= P(d_i|Frd, Wrk_i)P(Wrk_i|Frd) + P(d_i|Frd, Brk_i)P(Brk_i|Frd) \\ &= P(d_i|Frd, Wrk_i)P(Wrk_i) + P(d_i|Frd, Brk_i)P(Brk_i) \\ &= P(d_i|Frd, Wrk_i)p_i + P(d_i|Frd, Brk_i)(1-p_i) \end{aligned}$$

assuming the sensor working condition is independent of the nature of the target. These values are displayed in the columns $p(d_i|Frd)$ and $p(d_i|Foe)$ of Table IV. Their numerical values are in the present example:

$$\begin{aligned} P(d_i|Frd) &= 1 \times .99 + .5 \times .01 = .995 && \text{if } S_i \in HQS \text{ and } d_i = x_i \\ &= .5 \times .01 = .005 && \text{if } S_i \in HQS \text{ and } d_i = y_i \\ &= 1 \times .90 + .5 \times .10 = .95 && \text{if } S_i \in LQS \text{ and } d_i = x_i \\ &= .5 \times .10 = .05 && \text{if } S_i \in LQS \text{ and } d_i = y_i \end{aligned}$$

Given the observed data, we have:

$$P(Frd|data) \propto P(Frd) \times .995^{10} \times .0051 \times .951 \times .05^{18} = 1.72E - 26$$

Identically with foe, we get:

$$P(Foe|data) \propto P(Foe) \times .005^{10} \times .995^{1} \times .05^{1} \times .95^{18} = 1.93E - 25$$

Assuming equi prior probability of Friend and Foe ($P(Frd) = P(Foe) = .5$), we get:

$$P(Frd|data) = \frac{1.72E - 26}{1.72E - 26 + 1.93E - 25} = 0.08$$
$$P(Foe|data) = \frac{1.93E - 25}{1.72E - 26 + 1.93E - 25} = 0.92$$

Hence the probability analysis concludes that the target is a foe.

Strict Bayesian might argue that he $\alpha = .5$ was not correct and that another value for α must be used. This is not a real issue as once the α is determined, it is always possible to find a set of data so that the Bayesian and the TBM conclusions will diverge as strongly as here.

One might argue that this strict Bayesian analysis is not an adequate probability analysis, and that we should perform a sensitivity analysis, i.e., we must consider all possible values for α. The result becomes then totally uninformative and useless as we get $P(Frd|data)$ varying from 0 (when $\alpha = 1$) to 1 (when $\alpha = 0$), a truism of course.

C. TBM Analysis.

The TBM analysis leads to a conclusion opposite to the one reached by the strict Bayesian. It proceeds as follows. We must build the plausibility over $\{Frd, Foe\}$ given each type of observation and each sensor quality (see Table V). In the 'working' case, the sensors are perfect, hence the mass 1 on Frd with x_i and on Foe with y_i. When the sensor is 'broken', we are in a state of total ignorance about what might be the target, hence a mass 1 is given to $\{x_i, y_i\}$ for both possible data. Table V presents the corresponding plausibility functions on X_i (columns pl^{X_i}), and the values of $pl^{X_i}[Frd]$ and $pl^{X_i}[Foe]$ taking into consideration the p_i values (they are the weighted average of the former).

Target	Frd			Foe		
S_i Status	Wrk	Brk		Wrk	Brk	
$P(Status)$	p_i	$1-p_i$		p_i	$1-p_i$	
X_i	pl^{X_i}	pl^{X_i}	$pl^{X_i}[Frd]$	pl^{X_i}	pl^{X_i}	$pl^{X_i}[Foe]$
x_i	1	1	1	0	1	$1-p_i$
y_i	0	1	$1-p_i$	1	1	1

TABLE V

EXAMPLE 2. TBM APPROACH. FOR SENSOR S_i, VALUES OF $pl^{X_i}[T, S_i\ Status]$ WHERE THE TARGET T CAN BE Frd (FOR FRIEND) OR Foe AND S_i STATUS CAN BE Wrk (FOR WORKING) AND Brk (FOR BROKEN). p_i IS THE PROBABILITY SENSOR S_i STATUS IS Wrk. THE COLUMNS $pl^{X_i}[Frd]$ AND $pl^{X_i}[Foe]$ PRESENT THE PLAUSIBILITY ON X_i GIVEN THE TARGET IS Frd OR Foe, THUS THE LIKELIHOODS GIVEN TO THE TARGETS WHEN THE OBSERVED DATA IS x_i OR y_i, RESPECTIVELY.

In order to combine the data, we compute $pl^H[data] = \textcircled{\cap}_{1...30} pl^H[d_i]$. This is easily achieved using the commonality function $q^H[d_i]$ as:

$$q^H[data](T) = \prod_{i=1...30} q^H[d_i](T), \ \forall T \subseteq H.$$

Table VI presents the details of this computation and the resulting pignistic probabilities $BetP$. The TBM approach concludes that the target is a Friend with $BetP(Frd) = .91$.

Notice that if one had replaced the vector (1 1) of pl^{X_i} by (.5 .5) in the Brk columns of Table V, the results of the TBM analysis become the same is those of the probability approach. So the source of the difference between the results of the two approaches comes from the fact we represent total ignorance by equal probabilities in the probability

$q^H[x_i]$	HQS x_i	HQS y_i	LQS x_i	LQS y_i	$BetP$
$Friend$	1.00	0.01	1.00	0.10	.909
Foe	0.01	1.00	0.10	1.00	.091
H	0.01	0.01	0.10	0.10	

TABLE VI

EXAMPLE 2. INDIVIDUAL VALUES OF THE COMMONALITY FUNCTION $q^H[x_i]$. THE COLUMN $BetP$ PRESENTS THE PIGNISTIC PROBABILITIES OVER H.

approach and by a vacuous belief function in the TBM approach. Which representation is adequate is a matter of personal opinions.

D. Choosing between the two models.

As shown before, the two models, the Bayesian and the TBM, give totally opposite conclusions, which is quite disquieting. Its origin has been found in the representation of the uncertainty about how the sensor would react when broken.

We feel that we cannot leave each model stands against the other without giving readers some help to choose among them. A possible answer comes from the study of the informativity of the sources.

In the probabilistic setting, Shannon entropy has been defined in order to assess the informativity of distributions. The TBM can represent a broader range of uncertainty, Shannon entropy cannot be applied directly to bba's, and measures have been appropriately adapted to measure the informativity of a belief function [14].

According to these measures, the uniform distribution used in probability theory is not the most uninformative; it includes already 'something'. On the contrary, the vacuous belief is always the least informative. This might be used as an argument in favor of the TBM solution as it is based on less information than the probability solution.

VI. Computational Efficiency: Comparison with the Bayesian Model

In [4] the authors compare the Bayesian and the 'Dempster-Shafer' approaches in multisensor data fusion applied to a problem of target identification. They present a one and a two sensors example and a Monte Carlo simulation. They conclude that the use of belief functions is computationally less efficient than the use of probability functions. We repeat their study and show that in fact the converse holds.

Their approach is base on comparing the number of identical observations required by the two models so that the *a posteriori* probability of the actual hypothesis reaches a 0.99 threshold.

A. Example 3. One Sensor Problem: Mathematical Comparison

Let $H = \{h_1, h_2, \dots h_n\}$ denote a set of n hypotheses, for instance the type of aircraft under observation. One of these hypotheses, denoted h_0, corresponds to the actual one.

In order to determine the value of h_0, a sensor, like an Electronic Support Measure (ESM), makes an observation. Let X denote the set of possible values this observation can take. Suppose the sensor measurement is $x \in X$ and that its likelihoods are $l(h_1) = a$ and $l(h_i) = b$ for $i = 2, \dots, n$, $0 < b < a < 1$.

Suppose we collect k independent observations and each observation is the same $x \in X$. The likelihoods are then $l(h_1) = a^k$ and $l(h_i) = b^k$ for $i = 2, \dots, n$.

A.1 Probability Approach

The initial knowledge state is total ignorance, so we assume an *a priori* probability $P_0^H(h_i) = 1/n,\ \forall h_i \in H$.

With Bayes rule we compute the probability at step k. Let x^k represent the collected data, i.e., k times x. We have, $P^H[x^k](h_1) = a^k/(a^k + (n-1)b^k)$ and for $i = 2 \dots n$, $P^H[x^k](h_i) = b^k/(a^k + (n-1)b^k)$.

A.2 TBM Approach

For the TBM analysis, the GBT needs $pl^X[hi](x), i = 1 \dots n$ which are the likelihoods a^k and b^k of the previous section. To achieve the comparison, we need $BetP(h_1)$ after observing k times the same data x.

The bba on H is (see relation (11)):

$$m^H[x^k](h_1) = a^k(1-b^k)^{n-1}$$
$$m^H[x^k](h_1, h_2, \dots h_i) = a^k b^{k(i-1)}(1-b^k)^{n-i}$$
$$m^H[x^k](h_2, \dots h_i) = (1-a^k) b^{k(i-1)}(1-b^k)^{n-i}$$

or in general, $\forall h \subseteq H$

$$m^H[x^k](h) = a^{k|h_1 \cap h|}\ (1-a^k)^{1-|h_1\cap h|}\ b^{k|\overline{h_1}\cap h|}\ (1-b^k)^{(n-|\overline{h_1}\cap h|-1)}$$

A.3 Computing $BetP$

We derive the equation for $BetP^H[x^k]$ using the symbols as defined above.

Theorem VI.1:

$$BetP^H[x^k](h_1) = \frac{a^k}{nb^k}\ \frac{1-(1-b^k)^n}{1-(1-a^k)(1-b^k)^{n-1}} \tag{16}$$

Proof. Let $a^k = y, b^k = z$. By relation (10), we have:

$$\begin{aligned}
BetP^H[x^k](h_1) &= \sum_{h_1 \in h \subseteq H} \frac{m^H[x^k](h)}{|h|(1-m^H[x^k](\emptyset)} \\
&= \frac{1}{(1-m^H[x^k](\emptyset))} \sum_{h \subseteq \overline{h_1}} \frac{y}{1+|h|} z^{|h|}(1-z)^{n-1-|h|} \\
&= \frac{1}{(1-m^H[x^k](\emptyset))} \sum_{i=0}^{n-1} \frac{y}{1+i} \binom{n-1}{i} z^i (1-z)^{n-1-i} \\
&= \frac{y}{(1-m^H[x^k](\emptyset))} \frac{1}{nz}(1-(1-z)^n) \\
m^H[x^k](\emptyset)) &= (1-y)(1-z)^{n-1} \\
BetP^H[x^k](h_1) &= \frac{y}{nz}\ \frac{1-(1-z)^n}{1-(1-y)(1-z)^{n-1}}
\end{aligned}$$

□

A.4 Convergence Speed

We prove that the ratio $BetP^H[x^k](h_1)/P^H[x^k](h_1)$ is always larger than 1, thus that the TBM converges faster than the probability model, contrary to what [4] concludes.

Theorem VI.2: $BetP^H[x^k](h_1) \geq P^H[x^k](h_1)$.

Proof. Let $a^k = y, b^k = z$. By relation (16), we have:

$$\frac{BetP^H[x^k](h_1)}{P^H[x^k](h_1)} = \frac{y}{nz} \, \frac{1-(1-z)^n}{1-(1-y)(1-z)^{n-1}} \, \frac{y+(n-1)z}{y}$$

To prove the ratio is larger than 1, we must show that

$$(1-(1-z)^n)(y+(n-1)z) \geq nz(1-(1-y)(1-z)^{n-1})$$

which, after algebraic manipulations, becomes equal to:

$$(y-z)(1-(1-z)^n - nz(1-z)^{n-1}) \geq 0.$$

As $y \geq z$ by hypothesis, it is sufficient that:

$$1 \geq (1-z)^n + nz(1-z)^{n-1}.$$

This is equivalent to showing that:

$$\begin{aligned}
1 &\geq (1-z)^{n-1}(1+(n-1)z) \\
1-(1-z)^{n-1} &\geq (n-1)z(1-z)^{n-1} \\
\frac{1-(1-z)^{n-1}}{1-(1-z)} &\geq (n-1)(1-z)^{n-1} \text{ as } z \in [0,1) \\
\sum_{i=0}^{n-1}(1-z)^i &\geq (n-1)(1-z)^{n-1} \\
\sum_{i=0}^{n-1}(1-z)^{-(n-i-1)} &\geq n-1
\end{aligned}$$

As $z \in [0,1)$, $(1-z)^{-(n-i-1)} \geq 1$. Their sum is larger than n, what proves the inequality. $\square$

The ratio $BetP^H[d](h_1)/P^H[d](h_1)$ is thus always larger than 1 when $a > b$, for all $k = 1, 2 \ldots$. Therefore the number of steps k needed so that $BetP^H[d](h_1) \geq .99$ is never larger than the number of steps needed so that $P^H[d](h_1) \geq .99$.

This result contradicts the conclusions in [4]. The possible origin of the discrepancy is discussed in Section VI-D.

B. Example 4. Two Sensors Problem: Mathematical Comparison

Suppose we use two sensors S_1 and S_2 that observe the data x_1 and x_2, respectively. Let the likelihoods they generate in such cases be:

- $l_1(h_i) = a$ for $i = 1 \dots j$, and b for $i = j+1 \dots n$
- $l_2(h_i) = b$ for $i = 1 \dots j-1$, and a for $i = j \dots n$

where $a > b$. So alone, the sensors cannot discriminate the hypothesis h_j when it holds, whereas together, they do it nicely.

Suppose the same data x_1, x_2 has been collected k times, and the observations collected with the two sensors are independent, we can thus compute the likelihoods for the joint data. They are:

$$\begin{aligned} l_{12}(h_i) &= a^k b^k && \text{for } i = 1 \dots j-1 \\ &= a^{2k} && \text{for } i = j \\ &= a^k b^k && \text{for } i = j+1 \dots n \end{aligned}$$

We are thus back to the previous example, and the same proof shows that the pignistic probabilities computed in the TBM is always larger than the posterior probabilities computed in the probabilistic approach for any k.

C. Example 5. Monte Carlo Simulation

This part revisits the simulations used in [4]. We use the same data and compare the speed of convergence to the winning hypothesis by two decision systems: the classical Bayesian approach and TBM approach.

Contrary to the results presented in [4], the TBM-based algorithm is never slower than the Bayesian algorithm.

The experiment simulates a problem of decision concerning an aircraft engagement. The aircraft fled by the user detects another aircraft, and the question is to classify it. There are ten possible aircrafts. Table VII presents the various hypothesis, their class and their nature.

The user possesses a multisensor system to detect and recognize possible aircrafts: it consists of an Electronic Support Measure (ESM), an Identification Friend or Foe (IFF)

Identity	Type	Class	Nature
1	F15	Fighter	Friend
2	F16	Fighter	Friend
3	ATF	Fighter	Friend
4	B2	Bomber	Friend
5	Mig27	Fighter	Foe
6	Mig25	Fighter	Foe
7	Mig29	Fighter	Foe
8	Mig31	Fighter	Foe
9	Tu26	Bomber	Foe
10	Boeing	Commercial	Neutral

TABLE VII

THE VARIOUS HYPOTHESES OF THE PROBLEM.

and a Radar Sensors. We assume here that these three sensors have already been trained on the possible aircrafts. Tables VIII, IX, X present the confusion matrices, i.e., the conditional probabilities about the sensor observation for each possible aircraft.

The ESM sensor has been trained to discriminate between the 10 types. The IFF sensor can only discriminate between the two classes: Friend of Foe, and the Radar sensor can only distinguish between the three natures: Fighter, Bomber, or Commercial.

For the experiment, we varied three parameters. There was a possible misassociation: when detecting an aircraft, if a misassociation occurs, another aircraft is detected. The probability of a misassociation is 0.0 or 0.2 or 0.4. It was also possible that a sensor could not take a measure, sending in such a case an empty information. An empty information was modeled by a uniform distribution in the probability framework, and an vacuous belief function in the TBM framework. The probability that a sensor could not send a message ranges from 0 to 0.4 by 0.1 steps. The no-report problem could apply on each sensor separately, each pair of sensors and the three sensors simultaneously.

We rerun all the simulations performed in [4] and found out that the number of identical observations needed to reach a .95 *a posteriori* probability was always the same in both

Claimed to be	Actual aircraft									
	F15	F16	ATF	B2	Mig27	Mig25	Mig29	Mig31	Tu26	Boeing
F15	0.526	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053
F16	0.053	0.526	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053
ATF	0.053	0.053	0.526	0.053	0.053	0.053	0.053	0.053	0.053	0.053
B2	0.053	0.053	0.053	0.526	0.053	0.053	0.053	0.053	0.053	0.053
Mig27	0.053	0.053	0.053	0.053	0.526	0.053	0.053	0.053	0.053	0.053
Mig25	0.053	0.053	0.053	0.053	0.053	0.526	0.053	0.053	0.053	0.053
Mig29	0.053	0.053	0.053	0.053	0.053	0.053	0.526	0.053	0.053	0.053
Mig31	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.526	0.053	0.053
Tu26	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.526	0.053
Boeing	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.526

TABLE VIII

Confusion matrix for the ESM sensor.

Claimed to be	Actual aircraft									
	F15	F16	ATF	B2	Mig27	Mig25	Mig29	Mig31	Tu26	Boeing
Friend	0.909	0.909	0.909	0.909	0.091	0.091	0.091	0.091	0.091	0.091
Not Friend	0.091	0.091	0.091	0.091	0.909	0.909	0.909	0.909	0.909	0.909

TABLE IX

Confusion matrix for the IFF sensor.

approaches. We also discovered that the difference between the pignistic probability and the Bayesian probability observed when reaching the .95 threshold was positive in 70% with the pignistic probability, being larger than the Bayesian probability and within 1.0E-10 in 30% (what looks more like a rounding error). The difference was never negative. The largest difference encountered was 0.004. The TBM conclusions were always a little bolder than the one reached by the Bayesian approach.

These results indicate only that the claims made against the computational inefficiency of the belief based model are inexact.

D. Origin of the difference

There are three essentials differences between our GBT solution and the solution published in [32], [4], denoted hereafter the X solution.

- The interpretation of the confusion matrix data. In the X solution, the bba on H satisfies: $m^H[x](h_i) = l(h_i)$ and $m^H[x](H) = 1 - l(h_i)$ where $h_i \in H$ is the most likely hypothesis under x, i.e., $l(h_i) > l(h_j)\ \forall h_j \in H, h_j \neq h_i$. Its origin is not explained. Other

Claimed to be	Actual aircraft									
	F15	F16	ATF	B2	Mig27	Mig25	Mig29	Mig31	Tu26	Boeing
Fighter	0.833	0.833	0.833	0.083	0.833	0.833	0.833	0.833	0.083	0.083
Bomber	0.083	0.083	0.083	0.833	0.083	0.083	0.083	0.083	0.833	0.083
Commer.	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.833

TABLE X

Confusion matrix for the Radar sensor.

ad hoc suggestions can be found in [5]. In the GBT solution, we use $m^H[x](\overline{h_i}) = 1 - l(h_i)$ and $m^H[x](H) = l(h_i)$. The axiomatic justification of this choice can be found in [26].

• The relation used in the X solution to compute $m^H[x^k]$ after collecting k data (see ([32], eq. (7.28) page 247), and [4], eq. (1) page 570, where u_j must be read as u). The equation corresponds to Dempster's rule of combination if $m^H[x]$ is a simple support function, what happens for some of the cases analyzed by the X solution. Otherwise, the rule is ad hoc and does not correspond to Dempster's rule of combination.

To show the inadequacy of the published equation consider the case with $H = \{a, b\}$, $m^H[x](a) = m^H[x](b) = 1/4$ and $m^H[x](a, b) = 2/4$. Table XI presents the value of the normalized bba after collecting n times the same data x, with $n = 2$ and 3. The published equation gives for $n = 3$:

$$m^H[x3](a) = \frac{1/4\ [\ (1/4 + 2/4)2 + 2/4 + (2/4)2]}{2 * 1/4\ [\ (1/4 + 2/4)2 + 2/4 + (2/4)2] + (2/4)3} = \frac{21}{50}$$

whereas the solution obtained by Dempster's rule of combination is 19/46.

• The X solution focuses on comparing the *a posteriori* probabilities with the belief $bel^H[x_n](h_1)$ computed in the X model. Basing decision on bel is usually not advised. Had the authors used $pl^H[x_n](h_1)$, their conclusions would have been reversed. In the GBT, we use the pignistic probabilities for the comparisons, so comparing comparable objects. The justification for using $BetP$ and the inadequacy of bel or pl for decision making are presented in the overview [30].

These discrepancies lead to the published conclusions that we have shown to be incorrect. Study [4] fails as an argument against the computationally inefficiency of the TBM.

H	$m^H[x]$	$m^H[x2]$	$m^H[x3]$
$\{a\}$	1/4	5/14	19/46
$\{b\}$	1/4	5/14	19/46
$\{a,b\}$	2/4	4/14	8/46

TABLE XI

Computation of the normalized bba's $m^H[x^n] = \oplus_{i=1,\dots,n} m^H[x]$ using Dempster's rule of combination.

VII. Nested Sensors: Several Sensors per Level

In this example 6, we want to identify a vehicle. There are three types of vehicles: Buses ($h_1 = B$), Cars ($h_2 = C$) and Trucks ($h_3 = T$), two models of Buses, denoted h_{11}, h_{12}, two models of Cars, denoted h_{21}, h_{22} and two models of Trucks, denoted h_{31}, h_{32} (see Figure 1).

Suppose we have 4 sensors. The first, denoted S_0, measures the variable X_0 that can distinguish between Buses, Cars and Trucks. So its frame of discernment is $H_0 = \{h_1, h_2, h_3\}$. The second, denoted S_1, measures X_1 and can distinguish between the two models of Buses. The third, denoted S_2, measures X_2 and can distinguish between the two models of Cars. The fourth, denoted S_3, measures X_3 and can distinguish between the two models of Trucks. The frames of discernment of S_1, S_2, S_3 are H_1, H_2, H_3, respectively, where $H_i = \{h_{i1}, h_{i2}\}$. The granularity of each frame of discernment is important. So the frame of discernment H_0 has three singletons. Similarly the frame of discernment H_1 has two singletons, and the same holds for H_2 and H_3. The overall frame of discernment H has in fact six elements, the $h_{ij},\ i = 1,2,3,\ j = 1,2$. H_0 is a coarsening of H, whereas H_1, H_2 and H_3 are disjoint subsets of H.

The sensors S_0, S_1, S_2, S_3 produce the bba's $m^{H_i}, i = 0,1,2,3$, on their respective frames.

As defined H_1, H_2 and H_3 do not share a common refinement so conjunctive combination rules cannot be applied directly. In order to get a bba on H, we build the ballooning extension (see Section II-G) of the m^{H_i}'s on H. For instance $m^{H_1}(h_{11})$ will be extended on H so that it will be allocated to the $h_{11} \cup H_2 \cup H_3$. By construction, these extensions share the same frame of discernment and the combination rules can then be applied to

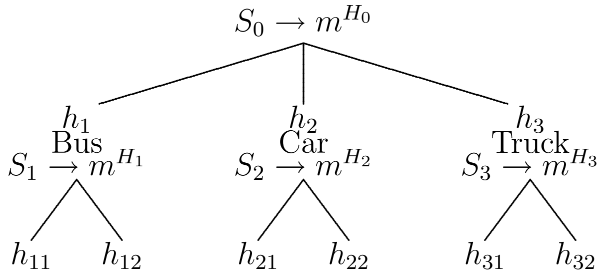

Fig. 1. Example 6. The tree of hypotheses and the domain of the four sensors with the beliefs they produce.

them.

When we state that S_1 can distinguish between the types of buses, we mean also that we have no idea whatsoever about how sensor S_1 would react if it was facing a car or a truck. This might occur if S_1 had never been used to measure the X_2 values on cars and X_3 values on trucks. Identical properties hold, up to permutation, for S_2 and S_3.

Suppose the four measurements are x_0, x_1, x_2, x_3. The four sets of likelihoods are presented in Tables XII and XIV.

	Bus h_1	Car h_2	Truck h_3
$l(h_i \mid x_0)$	0.60	0.40	0.10
$Bel^{H_0}[x_0]$	0.41	0.18	0.03
$Pl^{H_0}[x_0]$	0.77	0.51	0.13

TABLE XII

Example 6. Likelihoods on H_0 produced by the observations x_0, , and the normalized beliefs and plausibilities they induce on the singletons of H_0.

We first compute the bba $m^{H_0}[x_0]$ on H_0 given x_0 from the likelihoods produced by S_0 (i.e., the line $l(h_i|x_0)$ of Table XII). The resulting bba is presented in Table XIII. Its computation is done by applying equation (11). So for example

$$m^{H_0}[x_0](h_1) = .60 \times (1 - .40) \times (1 - .1) = .32$$

H_0	$\emptyset$	h_1	h_2	h_1, h_2	h_3	h_1, h_3	h_2, h_3	h_1, h_2, h_3
$m^{H_0}[x_0]$	0.22	0.32	0.14	0.22	0.02	0.04	0.02	0.02

TABLE XIII

Example 6. The basic belief assignment on H_0 given x_0, computed from the likelihoods of Table XII.

	Bus $i = 1$		Car $i = 2$		Truck $i = 3$	
	h_{11}	h_{12}	h_{21}	h_{22}	h_{31}	h_{32}
$l(h_{ij}\|x_i)$	0.45	0.55	0.90	0.10	0.40	0.70
$Bel^{H_i}[x_i]$	0.27	0.40	0.89	0.01	0.15	0.51
$Pl^{H_i}[x_i]$	0.60	0.73	0.99	0.11	0.49	0.85

TABLE XIV

Example 6. Likelihoods on H_1, H_2, H_3 produced by the observations x_1, x_2, x_3, , and the normalized beliefs and plausibilities they induce on the singletons.

For each sensor S_i, $i = 1, 2, 3$, we compute $m^{H_i \Uparrow H}[x_i]$, the ballooning extension of $m^{H_i}[x_i]$ on H.

We conjunctively combine these three bba and (the vacuous extension of) $m^{H_0}[x_0]$:

$$m^H[x_0, x_1, x_2, x_3] = (\circledcirc_{i=1,2,3}\, m^{H_i \Uparrow H}[x_i]) \circledcirc m^{H_0}[x_0]^{\uparrow H}.$$

This last bba is the final bba on H built from all collected data.

For practical applications, the computation can be tremendously speed up as, in practice, we hardly need all bbm, but only *bel* and *pl* (and maybe $BetP$) on the elements of H. Table XV presents these end results for these elements, i.e., the normalized Bel and Pl functions.

The analysis of the data show that S_0 supports that the target is a Bus (h_1), but after collecting all data, it appears that the h_{21} is the best supported target type (thus not a Bus). These data enhance the danger of premature decision making. Suppose we apply an iterated procedure by first observing S_0's data, and decide them to collect only S_1's data as far as the first step leads us to consider the target was a Bus. It would save the cost of collecting S_2 and S_3's data, but the end result would have been erroneous, as the h_{21} hypothesis would have been rejected, whereas it seems nevertheless the best hypothesis in the present case. This illustrates the dilemma between cost reduction obtained by taking intermediate decisions versus larger expenses resulting from delayed decisions with 'better' results.

In practical applications where cost reduction is an issue, a pre-posterior sensitivity analysis has to be realized at each step in order to decide if collecting further data would affect the results and are worth the effort. The method is essentially mimicking the strategy followed by the Bayesians. We do not explore this methodology further here, but it can 'easily' be performed within the TBM.

	h_{11}	h_{12}	h_{21}	h_{22}	h_{31}	h_{32}
$Bel^H[x_0, x_1, x_2, x_3]$	0.104	0.156	0.240	0.003	0.006	0.000
$Pl^H[x_0, x_1, x_2, x_3]$	0.397	0.485	0.530	0.059	0.059	0.103
$BetP^H[x_0, x_1, x_2, x_3]$	0.231	0.300	0.364	0.028	0.025	0.052

TABLE XV

Example 6. Normalized beliefs and plausibilities and pignistic probabilitys induced on the singletons by x_0, x_1, x_2, x_3.

VIII. Nested Sensors: One Sensor per Level

As in Section VII, in this example 7, we want to identify vehicles. They can be categorized in three types: Buses (h_1), Cars (h_2) and Trucks (h_3). Each type can be subdivided according to the auto-maker: Buses can be VanHool (h_{11}) or Mercedes (h_{12}), Cars can be VW (h_{21}), Audi (h_{22}) or Ford (h_{23}), Trucks can only be GMC (h_{31}). For each auto-maker, there are two models of vehicles (h_{ij1}, h_{ij2}), like 'Beetle' and Passat for VW, A4 and A6 for Audi... Figure 2 presents the tree describing the relation between the hypotheses.

There are three sensors, denoted S_1, S_2, S_3 which measure the values of the three variables X_1, X_2, X_3, respectively. S_1 is able to distinguish between the three types, S_2 between the 6 auto-makers, and S_3 between the 12 models. It means that:

- the frame of discernment $H = \{h_{ijk} : i = 1 \ldots 3, j = 1 \ldots n_i, k = 1, 2\}$ where h_{ijk} is one

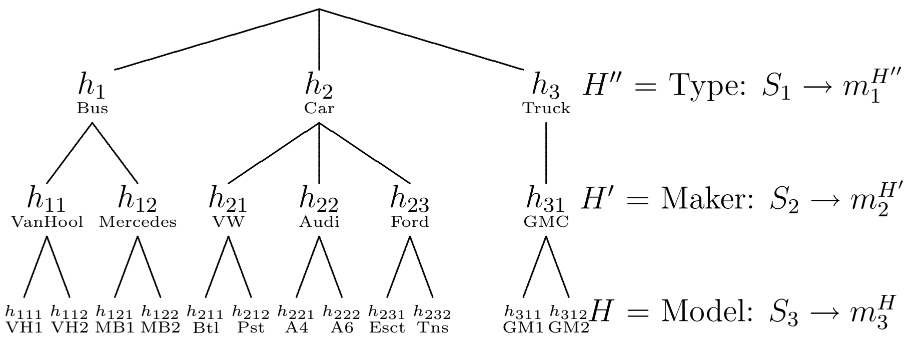

Fig. 2. Example 7. The tree of hypotheses and the domain of the three sensors with the beliefs they produce.

of the 12 models.

• there is a coarsening $H' = \{h_{ij} : i = 1, 2, 3, j = 1 \ldots n_i\}$ of H where the elements of H' are the 6 auto-makers.

• there is a coarsening $H'' = \{h_i : i = 1, 2, 3\}$ of H' where the elements of H'' are the 3 types.

• S_1 reports likelihoods on the types (H''), S_2 on the maker (H') and S_3 on the model (H). It results from the fact that, for S_1, we only know the probability over X_1 given h_1, h_2 and h_3. We have no information about what would be these probabilities given more refined hypothesis, like for instance h_{11}. So S_1 reports only likelihoods on H''. Similarly S_2 reports likelihoods on H', and S_3 on H.

• we denoted by $m_1^{H'}, m_2^{H'}$ and m_3^H the bba produced by the three sensors S_1, S_2, S_3, respectively.

The practical aim is to compute the normalized beliefs, plausibilities and pignistic probabilities for the singletons of H (the models) considering all available likelihoods. The computation could be done in a straightforward manner by vacuously extending $m_2^{H'}$ and $m_1^{H''}$ on H, and conjunctively combining the three bba's so defined on H. But this is not computationally efficient as it means we would have to work on the space 2^H. A more efficient algorithm can be described that needs only to produce the full bba on H'. As one can expect that the cardinality of the antepenultimate level in a tree is really smaller that the one of the last level, the computational benefit can be serious.

In the present example, suppose the three sensors have observed x_1, x_2, x_3, respectively. Let l denote the corresponding likelihoods. So we have three sets of likelihoods:

- $l_1(h_i)$ for $h_i \in H''$
- $l_2(h_{ij})$ for $h_{ij} \in H'$
- $l_3(h_{ijk})$ for $h_{ijk} \in H$

Let m_{123}^H denotes the bba obtained on H after conjunctively combining the bba's produced by the three sensors, and $m_{12}^{H'}$ the one on H' produced by combining sensors S_1 and S_2.

$$m_{12}^{H'} = m_2^{H'} \textcircled{\cap} m_1^{H'' \uparrow H'}$$
$$m_{123}^{H} = m_3^{H} \textcircled{\cap} m_{12}^{H' \uparrow H}$$

As we only need pl_{123}^H, bel_{123}^H on the singletons of H, very efficient algorithms can be used. Some are detailed in [6].

A. Computation

Tables XVI and XVII present some details of the computation for the case presented in Figure 2. The object is an VH1 (h_{111}). We use likelihoods of .1 for the wrong hypothesis, and .4 for the correct ones (see Table XVI). We do not detail the subsets of h_2 and h_3 as the values are all the same as those given in the columns h_2 and h_3. The terms $bel_1^{H''}$ and $pl_1^{H''}$ are computed from l_1 by equations (12) and (13). The bba $m_1^{H''}$ corresponding is detailed in Table XVII in the column $m_1^{H'' \uparrow H'}$. In that last table, we also presented the masses $m_2^{H'}$ computed from l_2, and the masses obtained by conjunctively combining $m_1^{H''}$ and $m_2^{H'}$ into $m_{12}^{H'}$ (the small masses are not presented). Table XVI presents then the beliefs and plausibilities on the singletons of H' and H. Finally, the normalized results on the singletons are displayed at the line Bel_{123}^H and Pl_{123}^H. The hypothesis VH1 (h_{111}) is strongly supported, VH2 (H_{112}) get a small support, and all other hypotheses can be neglected. We do not present the pignistic probabilities as they are always between Bel and Pl, and thus they would hardly bring any useful detail in the present context. Indeed $BetP_{123}^H(h_{111}) \geq .65$ whereas it will be less than .18 for all other hypotheses.

In conclusion, together the three sensors point strongly toward the fact the vehicle is a VH1. The purpose of this numerical example is only to illustrate the computation procedure and as such the data do not deserve a more detailed analysis.

H''	h_1				h_2	h_3	$m(\emptyset)$
l_1	.4				.1	.1	
$bel_1^{H''}$	.324				.054	.054	.486
$pl_1^{H''}$	.4				.1	.1	
H'	h_{11}		h_{12}		h_{2j}	h_{3j}	
l_2	.4		.1		.1	.1	
$bel_{12}^{H'}$	.139		.023		.006	.008	.786
$pl_{12}^{H'}$	.16		.04		.01	.01	
H	h_{111}	h_{112}	h_{121}	h_{122}	h_{2jk}	h_{3jk}	
l_3	.4	.1	.1	.1	.1	.1	
bel_{123}^{H}	.056	.009	.003	.003	.001	.001	.913
pl_{123}^{H}	.064	.016	.004	.004	.001	.001	
Bel_{123}^{H}	.65	.11	.03	.03	.01	.01	
Pl_{123}^{H}	.73	.18	.05	.05	.01	.01	

TABLE XVI

EXAMPLE 7. VALUE OF bel AND pl FOR THE SINGLETONS OF H'', H' AND H, AND THEIR NORMALIZED VALUES Bel AND Pl ON H. RIGHTMOST COLUMN GIVES THE MASSES GIVEN TO $\emptyset$ USED FOR NORMALIZATION.

B. Comparison with the likelihood solution

Reader may get the feeling that the solution provided by the TBM is not different from the one derived from the likelihood approach; in which case why to to use the TBM when the likelihood theory provided the same output in a simpler way. As shown in theorem III.1, this holds when the a priori on H is vacuous. It does not resist once *a priori* beliefs is introduced.

The next numerical example 8 presented in table XVIII uses a case with two sensors, one reporting on a frame H_1 with three elements $h_1, h_2, , h_3$, the second on H_2, a refinement of H_1 with five element $h_{11}, h_{12}, h_{21}, h_{22}, h_{31}$ where $h_{ij} \in h_i$ for all i and j. The table presents

h_{11}	h_{12}	h_{21}	h_{22}	h_{23}	h_{31}	$m_1^{H''\uparrow H'}$	$m_2^{H'}$	$m_{12}^{H'}$
0	0	0	0	0	0	.486	.354	.786
1	0	0	0	0	0		.236	.139
0	1	0	0	0	0		.039	.023
1	1	0	0	0	0	.324	.026	.015
0	0	1	0	0	0		.039	.006
0	0	0	1	0	0		.039	.006
0	0	0	0	1	0		.039	.006
0	0	0	0	0	1	.054	.039	.008
0	0	1	1	1	0	.054		
1	1	1	1	1	0	.036		
1	1	0	0	0	1	.036		
0	0	1	1	1	1	.006		
1	1	1	1	1	1	.004		
Total						1.0	.811	.989

TABLE XVII

Example 7. Subsets of H' with the bba's $m_1^{H''}$ extended on H', $m_2^{H'}$ and their conjunctive combination. The subsets with $m_{12}^{H'}$ small than .001 are omitted. The 6 leftmost columns represent subsets of H' with the 1 indicating which elements belong to the subset. At bottom line, sum of the masses displayed.

the likelihoods collected by sensor S_1 and S_2 that report on H_1 and H_2, respectively. The joint likelihoods on H_2 is obtained by multiplying the likelihoods in a pointwise manner. The most likely solution is h_{11}. The next line presents the pignistic probabilities $BetP_{12}$ which is maximal on h_{21}. The two methods disagree and the idea that the TBM reproduces the likelihood solutions is incorrect.

IX. Sensors on Partially Overlapping Frames

In this example 9, we present a case of non compatible frames and a method to extend the conjunctive combination rule in order to handle non compatible frames with some

H_1	h_1		h_2		h_3
l_1	.60		.75		.30
H_2	h_{11}	h_{12}	h_{21}	h_{22}	h_{31}
l_2	.60	.50	.45	.10	.10
l_{12}	**.360**	.300	.337	.075	.030
$BetP_{12}$	.321	.251	**.338**	.063	.026

TABLE XVIII

EXAMPLE 8. THE LIKELIHOODS l_1 ON SPACE H_1 AND l_2 ON H_2 A REFINEMENT OF H_1. AT BOTTOM, THE COMBINED LIKELIHOOD l_{12} AND THE PIGNISTIC PROBABILITIES $BetP_{12}$ ON H_2.

partial overlap. The solution is a 'careful' solution. This topic has been studied in [12], [13] who discuss the present careful solution but also present other bolder solutions.

Suppose a sensor S_1 that has been trained to recognize h_1 objects and h_2 objects and a second sensor S_2 that has been trained to recognize h_2 objects and h_3 objects (like h_1 = airplanes, h_2 = helicopters and h_3 = rockets). Sensor S_1 never saw an h_3 object, and we know nothing on how S_1 would react if it was observing an h_3 object. Beliefs provided by S_1 are always on the frame of discernment $\{h_1, h_2\}$. The same holds for S_2 with h_1 and h_3 permuted.

A new object X is presented to the two sensors. Both sensors S_1 and S_2 express their beliefs as $m_1^{H'}$ and $m_2^{H''}$, the first on the frame $H' = \{h_1, h_2\}$, the second on the frame $H'' = \{h_2, h_3\}$. How to combine these two bba's into a bba m_{12}^H on a common frame $H = \{h_1, h_2, h_3\}$?

The careful solution consists in applying the ballooning extension on the frame H to each bba and then conjunctively combining the results.

In this example the first sensor supports that X is h_1, whereas the second claims that X is h_2. If X had been h_2, how comes the first sensor did not say so? So the second sensor is probably facing an h_1 and just states h_2 because it does not know what an h_1 is. So we feel that the common sense solution is $X = h_1$, what is confirmed by $BetP_{12}^H$, the pignistic probability computed from m_{12}^H, as its largest value .655 is given to h_1. How probabilists would solve that problem without introducing artificial extra assumptions is not obvious.

H	$m_1^{H'}$	$m_2^{H''}$	$m_1^{H' \Uparrow H}$	$m_2^{H'' \Uparrow H}$	m_{12}^H	pl_{12}^H	$BetP_{12}^H$
h_1	.6				.42	.90	.655
h_2	.1	.7			.07	.32	.190
h_3		.2			.02	.30	.155
$\{h_1, h_2\}$	.3			.7	.21	.98	
$\{h_1, h_3\}$			.6	.2	.24	.93	
$\{h_2, h_3\}$		.1	.1		.01	.58	
$\{h_1, h_2, h_3\}$			.3	.1	.03	1	

TABLE XIX

Example 9. Basic belief assignment $m_1^{H'}$ and $m_2^{H''}$ on two partially overlapping frames, with their ballooning extensions on the common frame H and their conjunctive combination m_{12}^H on H with its related plausibility and pignistic probability functions.

X. Conclusions

The transferable belief model (TBM) is a model developed to represent quantified uncertainty based on belief functions. We have explained how to use the TBM for problems of target identification, these being considered as just a prototypical example of object classification. The major tools are the General Bayesian Theorem that permits to pass from the likelihoods to the posterior beliefs, and the pignistic transformation that permits the construction of the probabilities needed for decision making.

Our model mimics the probabilistic approach except that every probability function is replaced by a belief function. The latter being much more general than the former, we can handle degrees of uncertainty hard to represent in probability theory. In particular we can represent the state of total ignorance. It provides a solution to the problem of choosing the adequate prior in the diagnosis process. With the TBM, a prior representing total ignorance is available and can be used directly. Of course if justified priors are available, they are included in the model. The TBM reduces itself into the classical probability approach when all the ingredients needed for such an analysis are available.

We present an example where the TBM conclusion is diametrically opposed to the probability one, indicating thus that the choice of the model is a very serious issue.

Models based on belief functions have often been criticized for their computational inefficiency [32], [4]. This study has been used as an argument against the use of the TBM, so we reproduce it and find, to our own surprise, that the TBM approach is in fact computationally more efficient that its probability analogous.

We show then how to use the TBM for multisensor target identification problems. We consider cases of sensors collecting data sequentially on frames that are successively more and more refined. Incidentally we show that the results produced by the TBM are different from those derived from the likelihood approach.

We consider also a case where the likelihoods are known on strict and only partially overlapping subsets of the overall frame.

Our approach have been extended to other real life issues like 1) when the communication capacity of the sensors are limited and the sensor transmits just a few likelihoods, 2) when the sensors express the likelihoods on frames which granularity changes in an unknown way with each measure, or 3) when the sensor collected repeatedly the same type of information, hence data cannot be considered as distinct sources of evidence.

For the problem of multisensor data fusion, the TBM seems to offer a serious alternative to the probability model. To decide which model is the best is delicate as the term 'best' is hardly clearly defined. A nice property of the TBM is that it uses only the data really available and does not require the introduction of probabilities which values are unknown, or worse artificial, but necessary to apply the probability model.

Some very encouraging results have been obtained in real life contexts by [16], [17], [18] for a problem of multisensor an tipersonal mine detection. Other applications related to the use of the TBM for data fusion problems and detections have been developed in [15], [33], [2] and a very fruitful TBM based method for discriminant analysis has been introduced by Denœux and his group [8]. Ayoun and Smets [3] study the problem of the number of targets under observation, a problem that precedes the identification phase.

References

[1] A. Appriou. Probabilités et incertitude en fusion de données multi-senseurs. *Revue Scientifique et Technique de la Défense*, 11:27–40, 1991.

[2] A. Appriou. Multisensor signal processing in the framework of the theory of evidence. In *NATO/RTA, SCI Lecture Series 216 on Application of Mathematical Signal Processing Techniques to Mission Systems*, 1999.

[3] A. Ayoun and Ph. Smets. Data association in multi-target detection using the transferable belief model. *Inter. J. Intell. Systems*, 16:1167–1182, 2001.

[4] D. M. Buede and P. Girardi. A target identification comparison of Bayesian and Dempster-Shafer multisensor fusion. *IEEE Trans. SMC: A*, 27:569–577, 1997.

[5] F. Delmotte. Comparison of Bayesian and belief approaches in diagnostic. *Inter. J. Intell. Systems*, 16:963–981, 2001.

[6] F. Delmotte and P. Smets. Target identification based on the transferable belief model interpretation of Dempster-Shafer model. *IRIDIA-ULB Technical Report*, TR/IRIDIA/2003-06.

[7] A. P. Dempster. Upper and lower probabilities induced by a multiple valued mapping. *Ann. Math. Statistics*, 38:325–339, 1967.

[8] T. Denoeux. A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man and Cybernetics:*, 25:804–813, 1995.

[9] D. Dubois and H. Prade. The principle of minimum specificity as a basis for evidential reasoning. In B. Bouchon and R. R. Yager, editors, *Uncertainty in Knowledge-based Systems*, pages 75–84. Springer Verlag, Berlin, 1987.

[10] J. Gordon and E. H. Shortliffe. The Dempster-Shafer theory of evidence. In B. G. Buchanan and E. H. Shortliffe, editors, *Rule-Based Expert Systems: the MYCIN Experiments of the Stanford Heuristic Programming Project*, pages 272–292. Addison-Wesley, Reading, MA, 1984.

[11] Y. T. Hsia. Characterizing belief with minimum commitment. In IJCAI-91, editor, *Int. Joint Conf. on Artificial Intelligence*, pages 1184–1189. Morgan Kaufman, San Mateo, Ca, 1991.

[12] F. Janez and A. Appriou. Fusion of sources defined on different non-exhaustive frames. In IPMU-96, editor, *Information Processing and Management of Uncertainty*, 1996.

[13] F. Janez and A. Appriou. Théorie de l'évidence et cadres de discernements non exhaustifs. *Revue Traitement du Signal.*, 13, 1996.

[14] G. Klir and M. J. Wierman. *Uncertainty-based-information : elements of generalized information theory.* Physica-Verlag, Heidelberg, NY, 1998.

[15] G. Lohmann. An evidential reasoning approach to the classification of satellite images. In R. Kruse and P Siegel, editors, *Symbolic and Quantitative Approaches to Uncertainty*, pages 227–231. Springer-Verlag, Berlin, 1991.

[16] N. Milisavljevic, I. Bloch, and M. Acheroy. Characterization of mine detection sensors in terms of belief function and their fusion. In *Proc. 3rd Int. Conf. on Information Fusion (FUSION 2000)*, pages ThC3.15–ThC3.22, 2000.

[17] N. Milisavljevic, I. Bloch, and M. Acheroy. Modeling, combining and discounting mine detection sensors within Dempster-Shafer framework. In *Detection Technologies for Mines and Minelike Targets*, volume 4038, pages 1461–1472. SPIE Press, Orlando, USA, 2000.

[18] N. Milisavljevic, S. van den Broek, I. Bloch, P. B. W. Schwering, H. A. Lensen, and M. Acheroy. Comparison

of belief function and voting method for fusion of mine detection sensors. In *Detection and Remediation Technologies for Mines and Minelike Targets VI*, volume 4394. SPIE Press, Orlando, USA, 2001.

[19] A. Nifle and R. Raynaud. Un argument pour le choix entre décision pignistique et maximum de plausibilité en théorie de l'évidence. *GRETSI'97, Grenoble*, 1997.

[20] L. J. Savage. *Foundations of Statistics.* Wiley, New York, 1954.

[21] G. Shafer. *A Mathematical Theory of Evidence.* Princeton Univ. Press. Princeton, NJ, 1976.

[22] Ph. Smets. *Un modèle mathématico-statistique simulant le processus du diagnostic médical.* PhD thesis, Université Libre de Bruxelles, (available through University Microfilm International, 30-32 Mortimer street, London W1N 7RA, thesis 80-70,003), 1978.

[23] Ph. Smets. Bayes' theorem generalized for belief functions. In *Proc. ECAI-86*, pages 169–171, 1986.

[24] Ph. Smets. Belief functions. In Ph. Smets, E. H. Mamdani, D. Dubois, and H. Prade, editors, *Non Standard Logics for Automated Reasoning*, pages 253–286. Academic Press, London, 1988.

[25] Ph. Smets. Constructing the pignistic probability function in a context of uncertainty. In M. Henrion, R. D. Shachter, L. N. Kanal, and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence 5*, pages 29–40. North Holland, Amsterdam, 1990.

[26] Ph. Smets. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *Int. J. Approximate Reasoning*, 9:1–35, 1993.

[27] Ph. Smets. What is Dempster-Shafer's model? In R. R. Yager, J. Kacprzyk, and M. Fedrizzi, editors, *Advances in the Dempster-Shafer Theory of Evidence*, pages 5–34. Wiley, New York, 1994.

[28] Ph. Smets. The normative representation of quantified beliefs by belief functions. *Artificial Intelligence*, 92:229–242, 1997.

[29] Ph. Smets. The transferable belief model for quantified belief representation. In D. M. Gabbay and Ph. Smets, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, volume 1, pages 267–301. Kluwer, Doordrecht, The Netherlands, 1998.

[30] Ph. Smets. Decision making in a context where uncertainty is represented by belief functions. In R. P. Srivastava and T.J. Mock, editors, *Belief Functions in Business Decisions*, pages 17–61. Physica-Verlag, Heidelberg, Germany, 2002.

[31] Ph. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66:191–234, 1994.

[32] E. Waltz and J. Llinas. *Multisensor Data Fusion.* Artech House, Boston, 1990.

[33] L. M. Zouhal and T. Denœux. An adaptive k-NN rule based on Dempster-Shafer theory. In *Proc. of the 6th Int. Conf. on Computer Analysis of Images and Patterns (CAIP'95)*, pages 310–317, Prague, September 1995. Springer Verlag.